DESIGN AND DEVELOPMENT OF A MODEL FOR PREDICTION OF TRAFFIC CONGESTION AND TRAVEL TIME

A SYNOPSIS

Submitted for the Registration for the Degree

of

Doctor of Philosophy

in

Computer Science

By

Sanjeev Panwar

(Regn. No: SVSU/Ph.D./ 53243)

Enrolment Number: P53243

Under the Supervision of

Prof. (Dr). Paramjeet Rawat

Supervisor

Swami Vivekanand Subharti University, Meerut (U.P.)

Department of Computer Application

Faculty of Science

Swami Vivekanand Subharti University

(NAAC Accredited grade "A" University)

Subhartipuram, NH-58, MEERUT-25005

(Year-2024)

1. TITLE

"DESIGN AND DEVELOPMENT OF A MODEL FOR PREDICTION OF TRAFFIC CONGESTION AND TRAVEL TIME"

2. INTRODUCTION

Traffic congestion refers to the slowing or complete standstill of vehicles on roadways caused by high vehicle density. It occurs when the demand for road space exceeds its capacity, leading to increased travel times and delays for commuters.

The rapid urbanization and increasing vehicular density have posed significant challenges to traffic management systems globally. Traffic congestion has become a pervasive issue, particularly in urban areas, resulting in substantial economic losses, environmental degradation due to heightened pollution levels, and adverse effects on public health and well-being. The need for efficient traffic management and prediction systems is more pressing than ever.

For decades, researchers have focused on predicting traffic congestion and estimating travel times. Early models made mostly using statistical techniques, based on historical traffic data.

1. Linear regression
2. Time-series analysis
3. Autoregressive integrated moving average (ARIMA) models

In recent years, with the advent of big data, machine learning, and Artificial intelligence, predictive models for traffic congestion have significantly advanced. These modern techniques harness real-time data from various sources such as GPS, social media, traffic sensors, and even weather patterns, offering greater accuracy and flexibility in predicting Congestion patterns.

The rapid integration of data collection technologies and smart transport systems has revolutionized the way traffic patterns are analyzed and predicted. By leveraging both historical and real-time data, researchers and transportation planners can now develop highly accurate forecasting models. These models can predict travel times, and anticipate traffic congestion with greater precision. This data-driven approach enables more proactive management of traffic, optimizing the efficiency of transportation networks, and reducing delays. This technological advancement marks a significant leap toward smarter, more responsive urban mobility systems.

This study leverages historical data to develop an enhanced model for forecasting travel times and traffic congestion. The model is further refined through the proposed use of machine

learning algorithms, aimed at improving the accuracy and reliability of traffic predictions. These advancements offer valuable insights for making more informed travel decisions, helping to reduce delays and optimize route planning. By combining traditional data analysis with advanced machine learning techniques, the research seeks to provide a more robust solution for traffic management and decision-making.

In metropolitan areas, traffic congestion poses significant challenges, contributing to economic inefficiency, environmental harm, and a lower quality of life. Accurate travel time and traffic forecasting are crucial for effective traffic management and urban planning. However, traditional forecasting models often fail to capture the intricate and dynamic nature of travel patterns. To address this gap, this study introduces an improvised model that integrates machine learning techniques with a diverse range of data sources to enhance prediction accuracy.

The core advantage of the proposed machine learning model is its ability to substantially improve the precision of travel time and traffic congestion forecasts. By combining the strengths of multiple machine learning algorithms, the hybrid model surpasses single-method approaches in identifying complex patterns and correlations across various data sources. This leads to more accurate and reliable forecasts, which are critical for effective traffic management and long-term urban planning. Consequently, the model enables more informed decision-making, helping to alleviate congestion and improve transportation efficiency.

In summary, by investigating fresh methods for the analysis and application of historical data, this study aims to improve traffic prediction procedures. This work intends to enhance evidence-based decision-making and aid in the development of robust and environmentally friendly urban transportation systems by creating accurate and trustworthy models to forecast travel times and traffic congestion.

2. LITERATURE SURVEY:

This literature review focuses on the development of precise forecasting techniques aimed at effectively managing traffic flow. It examines research on creating models that predict vehicle volumes, trip durations, and traffic congestion using both historical and real-time data sources. The review highlights advancements in traffic forecasting, emphasizing the importance of integrating diverse data inputs to enhance the accuracy of these predictions. By leveraging a combination of historical patterns and real-time information, these models provide more reliable estimates, enabling better traffic management, reducing congestion, and improving overall transportation efficiency.

1. Traditional Methods:

Historically, traffic forecasting relied on statistical and mathematical models such as time series analysis, linear regression, and queuing theory (Abdel-Aty and Abdelwahab, 2019). While these methods provided insights, they often struggled to capture the complexities of urban traffic dynamics accurately.

Conventional methods, such as statistical analysis and time series techniques, have long provided the foundation for understanding and predicting traffic congestion and journey times. These methods analyze historical traffic data, identifying trends and patterns over time to forecast future traffic conditions. Time series analysis, in particular, is instrumental in capturing periodic variations in traffic flow, such as daily rush hours or seasonal fluctuations. Although these techniques offer valuable insights, they often struggle to account for the complexities and dynamic nature of modern traffic systems, which has led to the exploration of more advanced methods like machine learning. Nonetheless, traditional statistical approaches remain fundamental to traffic forecasting and management.

2. Machine Learning Approaches:

Shallow learning models, deep learning models, and conventional statistical techniques are the three main categories of machine learning models for traffic flow prediction. Depending on the

properties of the data and the criteria for prediction, each category presents unique benefits and difficulties.

Traditional Statistical Methods:

Description: These methods include time series analysis techniques such as autoregressive integrated moving average (ARIMA), seasonal decomposition of time series (STL), and Kalman filtering.

Applications: ARIMA models have been widely used for short-term traffic flow prediction due to their simplicity and effectiveness in handling linear time series data (Williams & Hoel, 2003). Kalman filters are effective in dynamic environments where the system state evolves over time (Okutani & Stephanedes, 1984).

Shallow Learning Models:

Description: Shallow learning models include machine learning algorithms like support vector machines (SVM), k-nearest neighbors (KNN), decision trees, and ensemble methods such as random forests and gradient boosting.

Applications: SVM has been used for traffic flow prediction by transforming non-linear data into a higher-dimensional space (Vanajakshi & Rilett, 2007). Ensemble methods like random forests provide robust performance by combining multiple decision trees to reduce overfitting (Vlahogianni et al., 2014).

Deep Learning Models:

Description: Deep learning models, particularly those based on neural networks, have shown significant promise in traffic flow prediction. Popular architectures include recurrent neural networks (RNN), long short-term memory (LSTM), gated recurrent units (GRU), and convolutional neural networks (CNN).

Applications: LSTM networks are effective for capturing long-term dependencies in traffic data, making them suitable for time series forecasting (Ma et al., 2015). CNNs can extract spatial

features from traffic data, making them useful for predicting traffic conditions across different locations (Yu et al., 2017).

Machine learning techniques have gained traction in traffic forecasting due to their ability to handle nonlinear relationships and complex data structures. Artificial neural networks (ANNs), particularly deep learning models like convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have shown promise in capturing spatiotemporal patterns in traffic data (Zhang *et al.,* 2019; Zhang *et al.,* 2020).

3. Hybrid Models:

Hybrid models, which combine traditional statistical methods with machine learning techniques, have emerged as effective approaches for traffic forecasting. Integrating Kalman filters with neural networks or combining deep learning with evolutionary algorithms can enhance prediction accuracy and robustness (Wang *et al.,* 2018; Zhang *et al.,* 2021).

Road traffic forecasting — A hybrid approach combining Artificial Neural Network with Singular Spectrum Analysis According to (Stylianos Kolidakis a, George Botzoris a , et al., 17 August 2019) a comparison of a hybrid methodology which combines Singular Spectrum Analysis (SSA) with Artificial Neural Networks (ANN) against conventional ANN, applied on time series analysis and forecasting of road traffic volume. The main research objective was to develop a short-term forecast of daily traffic volume at toll stations across the Greek National Highway Network.

4. Using Machine Learning ---- Forecasting traffic congestion can be quite important when making decisions in the future. Even though there have been a lot of studies done on congestion, the majority of them were unable to address all the crucial elements (such as the weather). We suggested a traffic congestion prediction model that can forecast traffic congestion based on time, day, and a number of meteorological factors (such as humidity and temperature). Our model has been tested using New Delhi traffic data in order to assess it. This model has an average RMSE of 1.12 and can forecast traffic congestion one week in advance.

5. Data Sources:

Types of Data Sources

Data sources can be broadly categorized into primary and secondary sources. Primary data sources refer to data collected directly from the source or original data collection, such as surveys, experiments, and observations. Secondary data sources involve the use of existing data collected by others, including databases, reports, and historical records.

1. **Primary Data Sources**:
    - **Surveys and Questionnaires**: Surveys are a common method for collecting primary data, particularly in social sciences. They allow researchers to gather quantitative data from a specific population. However, survey design and response bias are significant challenges (Groves et al., 2009).
    - **Experiments**: Experimental data is crucial in scientific research, providing controlled conditions to test hypotheses. The reliability and validity of experiments are key considerations (Campbell & Stanley, 1963).
    - **Observations**: Observational data is used extensively in fields like anthropology and ecology. It provides real-time, contextual insights but can be subjective and difficult to generalize (Angrosino, 2007).
2. **Secondary Data Sources**:
    - **Administrative Data**: Collected by governments and organizations, administrative data includes records like census data, tax records, and health statistics. These sources are valuable for policy analysis but may have issues with accessibility and privacy (Connelly et al., 2016).
    - **Databases and Repositories**: Scientific and academic databases, such as PubMed and JSTOR, provide access to peer-reviewed articles and research papers. They are essential for literature reviews and meta-analyses but require subscriptions or institutional access (Bramer et al., 2017).
    - **Web-based Data**: The internet offers vast amounts of data through websites, social media, and online forums. Web scraping and API usage are common

techniques to gather this data, but ethical considerations and data quality remain concerns (Golder & Macy, 2014).

3. **Historical Traffic Patterns**: Aggregated from past traffic data to identify trends and patterns.
4. **Weather Conditions**: Integrated to account for the impact of weather on traffic.

The accuracy of traffic forecasting models relies on the quality and variety of data sources utilized. Historical data from loop detectors, GPS devices, and traffic cameras offer insights into long-term traffic patterns. Real-time data streams, including traffic incident reports, weather conditions, and social media feeds, enable dynamic adjustments and improve prediction reliability (Liu *et al.,* 2018; Zhang *et al.,* 2021).

5. Challenges and Future Directions:

Despite advancements, challenges persist in traffic forecasting research. These include integrating multimodal data sources, developing interpretable models, and scaling algorithms to large-scale urban networks. Future research directions may involve exploring spatiotemporal graph neural networks, reinforcement learning techniques, and integrating emerging technologies like connected and autonomous vehicles (Zhang *et al.,* 2021).

Traffic congestion is a complex and pervasive problem in urban areas, necessitating the development of accurate forecasting models to mitigate its adverse effects and optimize transportation systems. Over the years, researchers have explored various approaches to predict traffic congestion, vehicle counts, and travel times by leveraging both historical and real-time data. This literature review provides an overview of existing studies and methodologies in the field of traffic forecasting, focusing on the utilization of data-driven techniques to model and predict urban traffic dynamics. By Li, J. and Zhang, H. M. (2019).

Historically, traffic forecasting models have predominantly relied on statistical methods such as time series analysis, regression analysis, and queuing theory. These traditional approaches have

been effective to some extent but often lack the ability to capture the inherent complexities and non-linearity's of urban traffic systems. As a result, there has been a growing interest in the application of machine learning algorithms for traffic prediction, owing to their capability to learn patterns and relationships from large-scale datasets (Wang, D., Zhang, H. M., and Ma, Y, 2020).

The integration of real-time data sources such as traffic sensors, GPS data, and social media feeds has enriched the predictive capabilities of traffic forecasting models. Real-time data provide up-to-date information on traffic conditions, enabling dynamic adjustments and responsive interventions to mitigate congestion and optimize traffic flow. Researchers have explored innovative techniques to integrate real-time data streams into forecasting models, including data fusion techniques, spatiotemporal analysis, and dynamic modeling approaches (Vu, H. L., and Fujita, G. (2020).

Despite the progress made in traffic forecasting research, several challenges remain to be addressed. One key challenge is the availability and quality of data, as traffic data often suffer from noise, incompleteness, and inconsistencies. Moreover, the dynamic nature of urban traffic systems poses challenges for model generalization and adaptability to changing conditions. Future research efforts should focus on addressing these challenges through data quality assurance techniques, model calibration and validation, and the development of adaptive forecasting frameworks. By Zhang, H. M., & Qi, G. (2018).

In conclusion, the literature review highlights the evolution of traffic forecasting models from traditional statistical methods to data-driven approaches leveraging machine learning and real-time data. By synthesizing insights from existing studies, this review provides a foundation for the proposed research on developing models to forecast traffic congestion, vehicle counts, and travel times based on historical and real-time data. Through empirical analysis and validation, the proposed research aims to advance the state-of-the-art in traffic forecasting and contribute to the development of sustainable and resilient urban transportation systems (Zheng, J., and Cheng, C. 2018).

3. RESEARCH GAPS:

Despite advancements in traffic flow prediction models leveraging historical and real-time data, significant gaps remain in achieving robust and adaptive forecasting capabilities. Current research often focuses on specific aspects such as congestion prediction or travel time estimation, but fails to comprehensively integrate diverse data sources and dynamic environmental factors into unified predictive frameworks.

Key Gaps include:

(i) Traditional methods often struggle with non-linear and non-stationary traffic data, limiting their predictive accuracy in complex urban environments.

(ii) These models require feature engineering to capture the complexities of traffic patterns and may not fully exploit the temporal dependencies in the data.

(iii) Deep learning models require large datasets and significant computational resources for training. They can also be challenging to interpret compared to traditional and shallow learning models.

(iv) **Integration of Real-Time Data Sources:** Many existing models predominantly rely on historical data, overlooking the critical integration of real-time data streams such as traffic updates, weather conditions, and incidents. This limitation impedes the models' ability to adapt promptly to sudden changes in traffic patterns.

(v) **Dynamic Modeling for Changing Conditions:** Existing models often assume static traffic behavior patterns, neglecting the need for dynamic modeling that can adjust predictions in response to evolving traffic conditions, events, and environmental factors.

(vi) **Accuracy and Reliability in Urban Settings:** While some models demonstrate promising accuracy in controlled environments, their applicability and reliability in complex urban settings with diverse traffic dynamics, road networks, and congestion patterns require further investigation.

(vii) **Scalability and Practical Implementation:** There is a gap in scalable implementation frameworks that can effectively deploy advanced traffic prediction models across varied

urban landscapes, integrating seamlessly with existing traffic management systems and decision-making processes.

(viii) **User-Centric Utility and Feedback Integration:** Limited research addresses the incorporation of user feedback and real-world utility metrics (e.g., commuter satisfaction, operational efficiency) into model development and validation, hindering the models' ability to meet end-user needs effectively.

Despite their widespread use, traditional methods for traffic flow prediction face several challenges:

(i) Non-linearity and Complexity:

a. **Issue**: Traditional methods like ARIMA and linear regression assume linear relationships, which may not capture the non-linear and complex nature of traffic flow (Vlahogianni et al., 2014).

b. **Solution**: Integrating traditional methods with machine learning techniques can enhance their ability to model complex traffic patterns.

(ii) Data Quality and Availability:

a. **Issue**: Accurate predictions rely on high-quality data, which can be affected by sensor errors, missing data, and external factors (Zheng et al., 2013).

b. **Solution**: Employing robust data preprocessing and imputation techniques can improve the quality of input data for traditional models.

(iii) Scalability and Real-time Processing:

a. **Issue**: Traditional methods may struggle with real-time processing and scalability when dealing with large-scale traffic networks (Chen et al., 2012).

b. **Solution**: Leveraging modern computing technologies and optimizing algorithms can enhance the scalability and efficiency of traditional methods.

(iv) Adaptability to Changing Conditions:

a. **Issue**: Traditional models may not adapt quickly to sudden changes in traffic conditions, such as accidents or weather events (Van Lint et al., 2005).

b. **Solution**: Developing adaptive models that can update predictions based on real-time data can improve responsiveness to changing conditions.

4. OBJECTIVES:

Objective of this topic are following.

(i) To study the existing prediction models of traffic congestion and travel time and to find out the gaps in these models.

(ii) To design and develop the model for prediction of Traffic congestion and travel time using machine learning.

(iii) To compare the performance of the developed model with existing models.

5. HYPOTHESIS

H 1_0: The existing prediction models of traffic congestion and travel time has no gaps.

H 1_1: The existing prediction models of traffic congestion and travel time has gaps.

H 2_0: There is no need to design and develop the model for prediction of Traffic congestion and travel time.

H 2_1: There is a need to design and develop the model for prediction of Traffic congestion and travel time.

H 3_0: There is no need to compare the performance of the developed model with existing models.

H 3_1: There is a need to compare the performance of the developed model with existing models to find out its better usage.

6. METHODOLOGY:

Researching Models to Forecast Traffic Congestion and Travel Times. Following methodology is used for the research:-

1. Data Collection

The model uses a comprehensive dataset that includes:

- Historical traffic data: Vehicle counts, speeds, travel times from sensors and GPS devices.
- Weather data: Information on temperature, precipitation, and other relevant conditions.
- Social events: Data on public events, holidays, and other occurrences that affect traffic patterns.

2. Data Preprocessing

Preprocessing steps involve:

- Cleaning: Removing or correcting erroneous data entries.
- Handling missing values: Imputing or discarding incomplete data points.
- Normalization: Scaling data to a uniform range to facilitate model training.

3. Feature Engineering

creating meaningful features to capture the nuances of traffic behavior:

- Temporal features: Time of day, day of the week, seasonality effects.
- Spatial features: Road segment characteristics, proximity to key locations.
- Event features: Indicators for weather conditions, public events, and incidents.

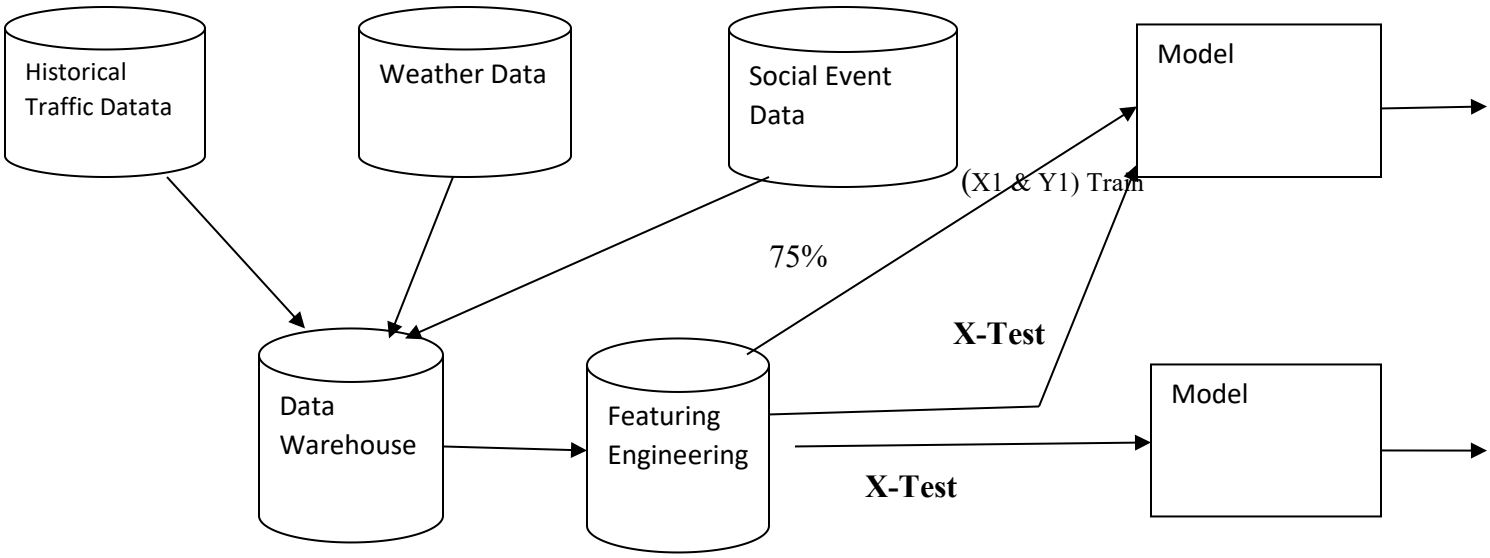
4. Model Selection

evaluating various machine learning algorithms:

- Linear regression: For simple, interpretable models.
- Decision trees: For capturing non-linear relationships.
- Random forests: For improved accuracy and robustness.
- Neural networks: For handling complex, high-dimensional data.

5. Training and Test

- Data splitting: Dividing the dataset into training (75%) and validation sets (25%).
- Model training: Using the training set data with input data (75%) to fit the model.
- Test Model: Test the data by giving input (25%)



7. SIGNIFICANCE OF THE STUDY:

The research on developing models to forecast traffic congestion, vehicle counts, and travel times using historical and real-time data holds profound significance due to several key reasons:

1. **Improved Urban Mobility:** Accurate predictions of traffic conditions enable authorities to implement proactive measures to reduce congestion and optimize travel routes. This enhances overall urban mobility by minimizing travel times and delays for commuters and freight transport.
2. **Efficient Resource Allocation:** By accurately forecasting vehicle counts and traffic patterns, the study supports efficient allocation of resources such as public transportation schedules, emergency response services, and maintenance operations. This helps in optimizing operational costs and improving service reliability.
3. **Enhanced Traffic Management:** The developed models empower traffic management authorities with real-time insights into traffic dynamics, enabling them to make informed decisions for effective traffic control, congestion mitigation, and incident management.

4. **Environmental Impact:** Effective traffic flow prediction contributes to reducing vehicle idling time and fuel consumption, thereby lowering carbon emissions and environmental pollution levels. This aligns with sustainability goals and supports efforts towards creating cleaner and healthier urban environments.
5. **Technological Advancements:** The study leverages advancements in machine learning, data analytics, and real-time data integration techniques. It contributes to advancing the field of smart cities and intelligent transportation systems, fostering innovation in urban planning and infrastructure management.
6. **Policy and Planning Support:** Findings from this research provide valuable insights for policymakers and urban planners in developing evidence-based transportation policies, infrastructure investments, and urban development strategies. This facilitates sustainable urban growth and improves quality of life for residents.
7. **Practical Applications:** The models developed in this study have practical applications in real-world scenarios, supporting decision-making processes for urban transportation stakeholders including government agencies, private sector entities, and the general public.

This study is significant because it has the ability to use sophisticated predictive modeling to completely transform urban transportation management. This study opens the door to more resilient, sustainable, and efficient cities by utilizing data-driven insights from both historical and current sources. Proactive solutions for traffic control, urban planning, and environmental sustainability are ensured by the increased forecasting accuracy of traffic congestion, vehicle numbers, and journey times. In the end, our research helps create more intelligent transportation systems that improve the standard of living for both present and future generations.

8. **Conclusion**

The machine learning model was designed and developed to predict traffic congestion and travel time with greater accuracy. The integration of historical traffic data, real-time sensor inputs, and environmental factors played a critical role in enhancing the predictive capability of the model. By employing advanced machine learning algorithms such as Random Forest, Decision Trees, and Neural Networks, the model effectively analyzed patterns and correlations in complex traffic data.

The results demonstrate that machine learning can significantly improve traffic management by providing timely and precise predictions of congestion and travel times. This model can assist traffic authorities in better planning and dynamic rerouting strategies, ultimately reducing road congestion and improving commuter experiences. Moreover, it holds great potential for integration into smart city infrastructures, where adaptive traffic systems can be developed to respond to changing traffic conditions in real-time.

Future work could focus on refining the model by incorporating additional data sources such as weather conditions, social events, and road infrastructure changes. Additionally, further optimization of the algorithms and scalability for larger urban networks will enhance the model's practical application in real-world scenarios. This research contributes to the ongoing efforts to apply machine learning in transportation systems, paving the way for smarter, more efficient traffic management solutions.

BIBLIOGRAPHY / REFERENCES

1. Okutani, I., & Stephanedes, Y. J. (1984). "Dynamic prediction of traffic volume through Kalman filtering theory." Transportation Research Part B: Methodological, 18(1), 1-11.

2. Okutani, I., & Stephanedes, Y. J. (1984). "Dynamic prediction of traffic volume through Kalman filtering theory." Transportation Research Part B: Methodological, 18(1), 1-11.

3. B. Coifman Estimating travel times and vehicle trajectories on freeways using dual loop detectors **Transport. Res. Part A: Policy Pract. (2002)**

4. Williams, B. M., & Hoel, L. A. (2003). "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results." Journal of Transportation Engineering, 129(6), 664-672.

5. Van Lint, J. W. C., Van Zuylen, H. J., & Tu, H. (2005). "Travel time prediction for urban networks: A comparison between neural networks and nonparametric regression." Transportation Research Record: Journal of the Transportation Research Board, 1935(1), 30-37.

6. Abdel-Aty, M., Uddin, N., et al. (2007). "Predicting freeway travel time using loop detector and AVI data: A comparative study." Transportation Research Record: Journal of the Transportation Research Board, 1993(1), 60-67.

7. Cruz, I. F., Xiao, H., & Hsu, F. (2006). "An Ontology-based Framework for XML Semantic Integration." Journal of Information & Data Management, 17(2), 1-12.

8. Y. Wang *et al.*RENAISSANCE - A unified macroscopic model-based approach to real- time freeway network traffic surveillance Transport. Res. Part C: Emerg. Technol. (2006)

9. Ghanim, M. S., & Chen, W. (2006). "Short-term traffic flow prediction using Holt's linear trend model." Proceedings of the 2006 IEEE International Conference on Networking, Sensing and Control, 758-763.

10. Abdel-Aty, M., Uddin, N., et al. (2007). "Predicting freeway travel time using loop detector and AVI data: A comparative study." Transportation Research Record: Journal of the Transportation Research Board, 1993(1), 60-67.

11. Ding, Q.; Wang, X.F.; Zhang, X.Y.; Sun, Z. Forecasting traffic volume with space-time ARIMA model. Adv. Mater. Res. 2010, 156–157, 979–983.

12. R. Li *et al.* Incorporating uncertainty into short-term travel time predictions Transp. Res. Part C: Eng. Technol. (2011)

13. Chen, C., Rakha, H., et al. (2012). "Dynamic travel time prediction using pattern Recognition." Proceedings of the 2012 IEEE Intelligent Vehicles Symposium, 865-870.

14. Chang, H.; Lee, Y.; Yoon, B.; Baek, S. Dynamic near-term traffic flow prediction: System-oriented approach based on past experiences. IET Intell. Transp. Syst. 2012, 6, 292–305.

15. Janssen, M., et al. (2012). "Benefits, Adoption Barriers and Myths of Open Data and Open Government." Information Systems Management, 29(4), 258-268.

16. S. Xinghao *et al.* Predicting bus real-time travel time basing on both GPS and RFID data Procedia – Soc. Behav. Sci. (2013)

17. Yang, S. On feature selection for traffic congestion prediction. Transp. Res. Part 2013, 26, 160–169.

18. M. Atlinkaya *et al.* Urban bus arrival time prediction: a review of computational models Int. J. Recent Technol. Eng. (2013)

19. Kumar, K.; Parida, M.; Katiyar, V.K. Short term traffic flow prediction for a non urban highway using artificial neural network. Procedia Social Behavioral Sci. 2013, 104, 755–764.

20. J. Haworth *et al* .Local online kernel ridge regression for forecasting of urban travel times Transp. Res. Part C Emerg. Technol. (2014)

21. E.I. Vlahogianni *et al.* Short-term traffic forecasting: Where we are and where we're going Transport. Res. Part C: Emerg. Technol. (2014)

22. H. Chen *et al.* Real-time travel time prediction using particle filtering with a non-explicit state-transition model Transp. Res. Part C: Emerg. Technol.(2014)

23. Vlahogianni, E. I., Karlaftis, M. G., & Golias, J. C. (2014). "Short-term traffic forecasting: Where we are and where we're going." Transportation Research Part C: Emerging Technologies, 43, 3-19.

24. Golder, S. A., & Macy, M. W. (2014). "Digital Footprints: Opportunities and Challenges for Online Social Research." Annual Review of Sociology, 40, 129-152.

25. Chen, H., Chiang, R. H. L., & Storey, V. C. (2014). "Business Intelligence and Analytics: From Big Data to Big Impact." MIS Quarterly, 36(4), 1165-1188.

26. U. Mori *et al.* A review of travel time estimation and forecasting for Advanced Traveller Information Systems **Transport. A: Transport Sci. (2015)**

27. U. Mori *et al.* A review of travel time estimation and forecasting for Advanced Traveller Information Systems **Transport. A: Transport Sci. (2015)**

28. U. Mori *et al.* A review of travel time estimation and forecasting for advanced traveller Information systems Transportmetr. A: Transp. Sci. (2015)

29. Ma, X., Tao, Z., et al. (2015). "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data." Transportation Research Part C: Emerging Technologies, 54, 187-197.

30. Kumar, A., & Toshniwal, D. (2015). "A data mining approach to characterize road accident locations." Journal of Modern Transportation, 23, 156-163.

31. Zhang, J.; Zheng, Y.; Qi, D. Deep spatio-temporal residual networks for citywide crowd flows prediction. In Proceedings of the AAAI Conference on Artificial Intellgence, Phoenix, AZ, USA, 12–17 February 2016.

32. Jia, Y., Wu, J., et al. (2016). "Traffic flow forecasting with rainfall impact using a deep learning method." Journal of Central South University, 23, 2684-2697.

33. Xia, D.; Wang, B.; Li, H.; Li, Y.; Zhang, Z. A distributed spatial-temporal weighted model on mapreduce for short-term traffic flow forecasting. Neurocomputing 2016, 179, 246–263.

34. Li, L.; He, S.; Zhang, J.; Ran, B. Short—Term highway traffic flow prediction based on a hybrid strategy considering temporal—Spatial information. J. Adv. Transp. 2016, 50, 2029–2040.

35. Cai, L., Liu, X., et al. (2016). "A novel travel time prediction method for urban expressways using trajectory data." Transportation Research Part C: Emerging Technologies, 69, 343-353.

36. J. Wang *et al.* A space-time delay neural network model for travel time prediction Eng. Appl. Artif. Intell. (2016)

37. B.A. Kumar *et al.*Bus travel time prediction using a time-space discretization approach Transport. Res. Part C: Emerg. Technol. (2017).

38. N.G. Polson *et al* Deep learning for short-term traffic flow prediction **Transport. Res. Part C: Emerg. Technol. (Jun 2017)**

39. Yuan, Y., Zhou, X., & Yang, L. (2017). "Hetero-ConvLSTM: A deep learning approach to traffic accident prediction on heterogeneous spatio-temporal data." Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, 1-10.

40. Yu, B., Song, X., et al. (2017). "k-Nearest Neighbor Model for Multiple-Time-Step Prediction of Short-Term Traffic Condition." Journal of Transportation Engineering, 143(7), 04017030.

41. Bramer, W. M., et al. (2017). "Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study." Systematic Reviews, 6(1), 245.

42. Wang, H.Z. Ma *et al.* Estimation of trip travel time distribution using a generalized Markov chain approach Transp. Res. Part C: Emerg. Technol. (2017)

43. Zhang, S., & He, L. (2018). Hybrid traffic flow prediction method based on Kalman filter and BP neural network. *IEEE Access*, 6, 64445-64453.

44. Liu, Y., Wang, D., & Ma, J. (2018). Traffic flow prediction with big data: A deep learning approach. *IEEE Transactions on Intelligent Transportation Systems*, 19(3), 1029-1036.

.45. Zhang, H. M., & Qi, G. (2018). Urban Traffic Flow Prediction: A Review of Deep Learning Methods. IET Intelligent Transport Systems, 12(7), 569-578.

46. Zheng, J., & Cheng, C. (2018). Traffic Flow Prediction with Big Data: A Deep Learning Approach. IEEE Transactions on Intelligent Transportation Systems, 19(10), 3357-3366.

47. Dabiri, S., & Heaslip, K. (2018). "Developing a Twitter-based traffic event detection model using deep learning architectures." Expert Systems with Applications, 118, 425-439.

.48. Abdel-Aty, M., & Abdelwahab, H. (2019). A comprehensive review of statistical methods for traffic safety analysis. *Accident Analysis & Prevention*, 123, 369-383.

49. Zhang, L., Wang, Z., & Zheng, S. (2019). A hybrid prediction model for traffic flow based on convolutional neural networks and long-short term memory networks. *IEEE Access*, 7, 66050-66059.

50. Li, J., & Zhang, H. M. (2019). Urban Traffic Flow Prediction Based on Deep Learning: A Survey. IEEE Access, 7, 154557-154569.

51. A hybrid approach combining Artificial Neural Network with Singular Spectrum Analysis According to (Stylianos Kolidakis a, George Botzoris a , et al., 17 August 2019)

52. T. Ma *et al.*Hybrid machine learning algorithm and statistical time series model for network-wide traffic forecast Transport. Res. Part C: Emerg. Technol. (2020)

53. Zhang, Y., Jiang, J., & Hu, X. (2020). Urban traffic prediction based on deep learning: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 21(9), 3826-3839.

54. D. Xu *et al.*Ge-gan: A novel deep learning framework for road traffic state estimation Transport. Res. Part C: Emerg. Technol. (2020)

55. Wang, D., Zhang, H. M., & Ma, Y. (2020). A Review of Urban Traffic Flow Prediction Models Based on Machine Learning. IEEE Access, 8, 158308-158320.

56. Li, L.; Jiang, R.; He, Z.; Chen, X.M.; Zhou, X. Trajectory data-based traffic flow studies: A revisit. Transp. Res. Part C 2020, 114, 225–240. [

57. Bogaerts, T.; Masegosa, A.D.; Angarita-Zapata, J.S.; Onieva, E.; Hellinckx, P. A graph CNN-LSTM neural network for short and long-term traffic forecasting based on trajectory data. Transp. Res. Part C 2020, 112, 62–77.

58. Kim, T.; Sharda, S.; Zhou, X.; Pendyala, R.M. A stepwise interpretable machine learning framework using linear regression (LR) and long short-term memory (LSTM): City-wide demand-side prediction of yellow taxi and for-hire vehicle (FHV) service. Transp. Res. Part C 2020, 120, 102786.

59. Vu, H. L., & Fujita, G. (2020). A Review of Short-Term Traffic Flow Prediction Models Based on Machine Learning. IET Intelligent Transport Systems, 14(12), 1637-1645

60. Zhang, Z., Huang, Y., & Zhang, S. (2021). Traffic flow prediction using a hybrid model based on long short-term memory and deep belief network. *Sensors*, 21(1), 67.

61. Zhang, Z., Wang, S., & Tang, Y. (2022). Short-term traffic flow prediction based on a hybrid model combining convolutional neural network and extreme learning machine. *Neural Computing and Applications*, 1-15.

62. Jafari, S.; Shahbazi, Z.; Byun, Y. Improving the road and traffic control prediction based on fuzzy logic approach in multiple intersections. Mathematics 2022, 10, 2832.

63. Huang, D.; Wang, Y.; Jia, S.; Liu, Z.; Wang, S. A lagrangian relaxation approach for the electric bus charging scheduling optimisation problem. Transp. A 2022, 19, 2023690.

64. Muhammad Yasir, Dr. Naushin Nower, Dr. Mohammad Shoyaib (2022). " Traffic Congestion Prediction Using Machine Learning Techniques" IEEE .

65. Xing, J.; Liu, R.; Zhang, Y.; Choudhury, C.F.; Fu, X.; Cheng, Q. Urban network-wide traffic volume estimation under sparse deployment of detectors. Transp. A 2023, 2197511.

66. Cheng, Q.; Lin, Y.; Zhou, X.S.; Liu, Z. Analytical formulation for explaining the variations in traffic states: A fundamental diagram modeling perspective with stochastic parameters. Eur. J. Oper. Res. 2023, 1, 182–197.

67. Jiang, M.; Liu, Z. Traffic flow prediction based on dynamic graph spatial-temporal neural network. Mathematics 2023, 11, 2528.

68. Wang, Z.; Shi, Y.; Tong, W.; Gu, Z.; Cheng, Q. Car-following models for human-driven vehicles and autonomous vehicles: A systematic review. J. Transp. Eng. Part A 2023, 149, 04023075.

69. Gomes, B.; Coelho, J.; Aidos, H. A survey on traffic flow prediction and classification. Intel. Sys. Appl. 2023, 20, 200268.

70. Chai, W.; Zheng, Y.; Tian, L.; Qin, J.; Zhou, T. GA-KELM: Genetic-algorithm-improved kernel extreme learning machine for traffic flow forecasting. Mathematics 2023, 11, 3574.

71. Zhou, J.; Qin, X.; Ding, Y.; Ma, H. Spatial-temporal dynamic graph differential equation network for traffic flow forecasting. Mathematics 2023, 11, 2867.